A fast and optimal multi-frame blind deconvolution algorithm for high-resolution ground-based imaging of space objects--Journal Article (Preprint)

Charles L. Matson et al.

1 April 2008

Technical Paper

AIR FORCE RESEARCH LABORATORY
Directed Energy Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-04-2008	Technical Paper	Jan 1, 2007- Apr, 1 2008

4. TITLE AND SUBTITLE

A fast and optimal multi-frame blind deconvolution algorithm for high-resolution ground-based imaging of space objects--Journal Article (Preprint)

5a. CONTRACT NUMBER
In- House (DF701944)

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)

Charles L. Matson, Kathy Borelli, Stuart Jefferies, Charles C. Beckner, Jr., E. Keith Hege, and Michael Lloyd-Hart

5d. PROJECT NUMBER
2311

5e. TASK NUMBER
S3

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

AFRL/RDSA
3550 Aberdeen Ave. SE
Kirtland AFB, NM 87117-5776

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
3550 Aberdeen Ave. SE
Kirtland AFB, NM 87117-5776

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RDSA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-RD-PS-TP-2008-1007

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited.

13. SUPPLEMENTARY NOTES
Submitted for publication in The Journal of Applied Optics, January 2009. "GOVERNMENT PURPOSE RIGHTS."

14. ABSTRACT
We report on a multi-frame blind deconvolution algorithm that we have developed for imaging through the atmosphere. The algorithm has been parallelized to a significant degree for execution on high-performance computer, with an emphasis on distributed-memory systems so that it can be hosted on commodity clusters. As a result, image restorations can be obtained in seconds to minutes. We have compared the quality of its output to the associated Cramer-Rao lower bounds and demonstrated that, for the cases we have tested, it achieves or closely approaches these bounds. In this paper, we describe the algorithm and its parallelization in detail, demonstrate the scalability of its parallelization across associated Cramer-Rao lower bounds, and present image restorations obtained using data collected with ground-based telescopes.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			SAR	47	Charles Matson
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified			505-846-2049

A fast and optimal multi-frame blind deconvolution algorithm for high-resolution ground-based imaging of space objects

Charles L. Matson,[1,*] Kathy Borelli,[2] Stuart Jefferies,[3] Charles C. Beckner, Jr.,[1] E. Keith Hege,[4]

and Michael Lloyd-Hart[4]

[1]*Air Force Research Laboratory, 3550 Aberdeen Ave SE, Kirtland AFB, NM*

[2]*KJS Consulting, Haiku, HI*

[3]*Institute for Astronomy, University of Hawaii, Kula, HI*

[4]*The College of Optical Sciences, University of Arizona, Tucson, AZ*

[*]*Corresponding author: charles.matson@kirtland.af.mil*

ABSTRACT

We report on a multi-frame blind deconvolution algorithm that we have developed for imaging through the atmosphere. The algorithm has been parallelized to a significant degree for execution on high-performance computers, with an emphasis on distributed-memory systems so that it can be hosted on commodity clusters. As a result, image restorations can be obtained in seconds to minutes. We have compared the quality of its output to the associated Cramér-Rao lower bounds and demonstrated that, for the cases we have tested, it achieves or closely approaches these bounds. In this paper, we describe the algorithm and its parallelization in detail, demonstrate the scalability of its parallelization across distributed-memory computer nodes, show the results of comparing sample variances of its output to the associated Cramér-Rao lower bounds, and present image restorations obtained using data collected with ground-based telescopes.

OCIS codes: 100.1455, 100.3020, 110.3055, 110.4155

1. INTRODUCTION

The term "blind deconvolution" refers to methods that jointly estimate an object and a blurring function from a measurement that consists of their convolution (usually corrupted by noise as well). It is straightforward to show that the blind deconvolution problem does not have a unique solution without including additional information in the estimation process. The additional information that can or should be included depends on the physics of the measurement and the available a priori knowledge about the object and the blurring function. There exists a plethora of blind deconvolution publications describing algorithms and/or theoretical results, differing primarily in the additional information assumed to be known and the means by which the object and the blurring function are estimated. A review article and its companion, published in 1996, provide a nice taxonomy of the types of additional knowledge commonly used and the algorithms employed to carry out the estimation process [1,2]. Although many years have passed since the articles were published, they are still excellent reviews for those unfamiliar with blind deconvolution.

The use of the term "blind deconvolution" occurred first in Ref. 3 to describe methods for restoring images and audio recordings. These methods require a separate set of measurements of so-called prototype objects that share statistical similarities with the object of interest in the measurements in order to recover both the blurring function and the object. Although these methods generalized the standard deconvolution problem (where the blurring function is assumed to be known), the requirement to have a set of prototype images limited the applicability of the methods. A number of years later, it was shown that the blind deconvolution problem for signals of dimension greater than one is uniquely solvable (up to a scale factor and a spatial shift) using only the information that the object and the blurring function have finite supports [4]. This result eliminated the need for a separate set of prototype images.

One important application of finite-support-based blind deconvolution is high-resolution astronomical imaging using ground-based telescopes, for at least two reasons. The first reason is that astronomical objects have finite size and are on a black background, and thus have finite support. The second reason is that the structure of the blurring functions in short-exposure atmospherically-blurred

images is such that the deconvolution operation can bring about significant increases in resolution. The blurring function, or point spread function (PSF), is dominated by atmospheric turbulence effects. Atmospheric turbulence blurring is the result of spatial and temporal fluctuations in the index of refraction of the atmosphere. In long-exposure images (integration times of seconds or longer), multiple atmospheric turbulence realizations combine to produce a long-exposure PSF that is essentially bandlimited to angular frequencies less than r_o/λ, where λ is the observation wavelength and r_o is the Fried parameter [5] that has a nominal value of 10-20 cm for visible wavelengths at astronomical telescope sites. Because of the structure of this PSF, deconvolution methods produce only nominal improvements in resolution. Furthermore, since the diffraction-limited resolution of telescopes is D/λ, where D is the diameter of the primary mirror of the telescope, and since modern astronomical telescopes have diameters of 8 meters or more, long-exposure PSFs dramatically reduce the theoretical resolution limits of telescopes.

The situation is significantly different for short-exposure PSFs (where the atmospheric turbulence remains fixed during the exposure time) because they have a structure that permits deconvolution-based resolution increases of up to a factor of D/r_o more than is present in long-exposure images. The first published deconvolution method to exploit this structure of short-exposure PSFs (Labeyrie's technique) generated deconvolved energy spectra with diffraction-limited resolution [6]. Labeyrie's technique was then extended to other techniques (such as cross-spectrum [7] and bispectrum [8] estimation) that permit the recovery of Fourier phase spectra as well as energy spectra so that diffraction-limited images can be obtained [9]. All of these techniques (called speckle imaging techniques) require as input a set of short-exposure images of an object where the object is common to all the short exposures but the atmospheric blurring is different for each image. They then estimate average values of the above-mentioned quantities for which diffraction-limited information is retained in the averaging process. The object Fourier phase can be extracted from the average image cross spectrum or bispectrum without deconvolution, but the image energy spectrum is equal to the object energy spectrum multiplied by the atmosphere energy

spectrum. Thus it is necessary to collect a separate set of measurements of an unresolved star to generate an atmosphere energy spectrum estimate to be used for deconvolution. For this reason, speckle imaging techniques have much in common with the techniques first referred to as blind deconvolution techniques [3], where the separate star measurements are the equivalent of the prototype image measurements.

In the field of astronomical imaging, the term "blind deconvolution" refers to deconvolution techniques that do not require separate measurements to determine the atmospheric PSF. The first blind deconvolution algorithm applied to atmospherically-blurred images, to the authors' knowledge, was the seminal work of Ayers and Dainty [10], where they used an iterative implementation reminiscent of the Gerchber /Saxton phase retrieval algorithm [11] and enforced a positivity constraint as well as a support constraint. The Ayers/Dainty algorithm was quickly followed by a version that works with complex as well as real images [12]. Because the convergence and noise properties of the Ayers/Dainty algorithm are not optimal, many other blind deconvolution implementations for removing atmospheric turbulence blurring from image data alone have been proposed, including projections on convex sets [13], neural networks [14], maximum-likelihood estimation (MLE) [15], non-negative matrix factorization [16], simulated annealing [17], and maximum a prior estimation [18]. It was soon realized that additional datatypes that were collected simultaneously with the image-plane data can produce higher-quality restorations. Such datatypes include a separate image-plane measurements that have purposely-induced phase distortions such as defocus [19], wavefront-sensor data [18], and imaging at multiple wavelengths [20].

A particularly-fruitful approach to blind deconvolution is to use a set of short-exposure images (measurement frames) of an object where the object is common to all the images but the blurring is different for each image. This type of blind deconvolution is commonly called multi-frame blind deconvolution (MFBD). The usual way to generate this set of images is to collect a temporal series of short-exposure images of the object. The first use of the MFBD term was in the context of jointly estimating both the unblurred object and the PSFs for each of the measurement frames [21], but this was not the first approach to using this type of data. The first approach was in the context of speckle imaging,

4

where the energy spectrum of the system was calculated from the measurement frames instead of using a separate set of measurements of an unresolved star [22]. Shortly thereafter, an approach was published that used a set of short-exposure images sequentially in a modified Ayers/Dainty algorithm [23], and then extended to use the short-exposure images jointly [24]. A key benefit of the approach described in Ref. 21 is that a penalized maximum-likelihood algorithm is used, which tends to be more robust than Gerchber/Saxton algorithms or projections onto convex sets.

The vast majority of the research in blind deconvolution is concentrated in four different areas: (1) algorithm development, (2) types of additional knowledge to include in algorithms, (3) theoretical results on speed and convergence, and (4) applications to real data. In addition to these areas, another important area of blind deconvolution research is the development of algorithm-independent image-quality performance bounds such as Cramér-Rao lower bounds (CRBs) [25]. A number of papers have been published including the use of CRBs to compare two algorithms assuming white Gaussian noise and single-input single-output systems [26], asymptotic CRB expressions for a multiple-input multiple-output system with white non-Gaussian noise [27,28] and analyses of phase-diversity-based blind deconvolution [29,30].

A second important area of blind deconvolution research for which few published results are available is investigating algorithm architectures that take advantage of parallel computational resources and demonstrations of improved performance. One published approach is based on segmenting the blurred images into a number of smaller images, each of which can be processed in parallel [14]. The authors do not discuss the speed improvements achieved by this type of parallelization in any detail. A second approach to blind and parallel processing is to set up the conjugate gradient search algorithm and the filter bank structure to be in parallel [31]. Again, no speed improvements are discussed. A third approach is based upon specialized hardware; however, no discussion of how to implement blind deconvolution on this architecture is given [32].

In this paper, we describe our MFBD algorithm that we call the physically-constrained iterative deconvolution (PCID) algorithm. It is a MLE-based algorithm that utilizes a conjugate-gradient

minimization routine and regularization with the use of either a penalty term or a Fourier-domain filter. Because our application is astronomical imaging, we have little or no a priori information about the objects to be imaged other than generic constraints such as support and positivity constraints; therefore, the object is always estimated pixel-by-pixel in the image domain. Because we do assume that the system PSF results from atmospheric turbulence and telescope diffraction, we have included the ability to estimate the PSFs either pixel-by-pixel in the image domain or in terms of a Zernike-based expansion of the phase in the pupil of the telescope [9]. We also describe how we parallelized PCID to run on distributed-memory parallel computers in a manner that scales well for processor counts of one hundred or more, depending upon the speed of the communication network between the nodes, the number of pixels in a measurement frame, and the number of measurement frames used in a single image restoration. We show that image restorations can be generated in seconds to minutes for using a Cray XD1 supercomputer [33]. We demonstrate that the variances of PCID-generated object estimates achieve or closely approach their CRBs for a variety of algorithm parameters.

The outline of the paper is as follows: in Section 2 we describe all but the parallelized architecture of the PCID algorithm, in Section 3 we describe the parallel architecture of PCID and demonstrate PCID's performance as a function of the number of processors used by the algorithm, in Section 4 we present CRB theory as it applies to PCID and compare PCID performance to these CRBs, in Section 5 we give several examples of PCID restorations, and in Section 6 we give conclusions and anticipated future work.

2. PCID ALGORITHM DESCRIPTION

In this section we first present the theory underlying the PCID algorithm and then describe all the implementation issues associated with producing the desired estimates. The PCID algorithm incorporates features that increase the probability and speed of finding the global minimum, achieve superresolution by de-aliasing, handle non-idealities such as clipped images due to jitter and/or object size relative to the detector size, and generate many of the PCID inputs automatically.

6

A. PCID theory

The forward model used in the PCID algorithm is the standard linear imaging model given by

$$i_m(x) = o(x) * h_m(x) + n_m(x),$$

(1)

where $i_m(x)$ is the mth measurement frame, $o(x)$ is the true object, $h_m(x)$ is the mth PSF, $n_m(x)$ is the mth noise realization, x is a two-dimensional spatial location vector, bold-face type indicates vector and matrix quantities, and $*$ denotes convolution. The noise term incorporates both Gaussian and Poisson noises that are always present in digital images collected using digital detectors such as CCD cameras. The assumption that a noise-free image is the convolution of the true object and the PSF means that the system is assumed to be space invariant. For many applications, this is a valid assumption; however, even when the system violates the space-invariant assumption, this model still produces excellent image restorations as will be demonstrated in Section 5. We call the PSF invertible if its Fourier transform contains no zeros, and non-invertible otherwise.

The object and PSF estimates are generated by minimizing a cost function that is based on MLE theory [34]. With the use of the central limit theorem, valid for all but quite dim objects, we model $n_m(x)$ as a spatially-independent, zero-mean Gaussian random process with a variance equal to $\sigma_m^2(x) + i_m(x)$, where $\sigma_m^2(x)$ is the variance of the Gaussian noise and $i_m(x)$ is the variance of the Poisson noise in the mth frame. For the Gaussian noise model, the general form of the MLE-based cost function $J\left[\hat{o}(x), \hat{h}_1(x), ..., \hat{h}_M(x)\right]$ that is minimized to generate estimates of the object and the PSFs is given by

$$J\left[\hat{o}(x), \hat{h}_1(x), ..., \hat{h}_M(x)\right] = \sum_{m=1}^{M} \sum_{n=1}^{N^2} \frac{1}{\sigma_m^2(x_n) + i_m(x_n)} \left[i_m(x_n) - \hat{i}_m(x_n)\right]^2 + \gamma g\left[\hat{o}(x_n), \hat{h}_1(x_n), ..., \hat{h}_M(x_n)\right],$$

(2)

7

where M is the number of measurement frames, N^2 is the number of pixels in each measurement frame, $\hat{o}(x), \hat{h}_1(x),...,\hat{h}_M(x)$, are the estimated object and PSFs, $\hat{i}_m(x)$ is the estimated m^{th} measurement frame, $g()$ is a function used for regularization, $\{x_n\}$ are discrete spatial locations corresponding to pixels in a detector, and γ is the Lagrange multiplier weighting the relative importance of the regularization term and the data-matching term. In the PCID implementation of Eq.(2), the photon noise weighting can be based on the measurement as shown, or can be based on the estimate of the measurement. In the latter case, $i_m(x)$ is replaced with $\hat{i}_m(x)$.

B. PCID implementation

All the implementation issues for the PCID algorithm are described in this section. These issues are: (1) how $\hat{o}(x)$ and $\hat{h}_m(x)$ are parameterized, (2) how positivity and support are enforced, (3) how regularization is implemented, (4) the means that we use to minimize $J[\hat{o}(x), \hat{h}_1(x),...,\hat{h}_M(x)]$, (5) the method to achieve de-aliasing-based superresolution, (6) two uniqueness issues, and (7) some algorithm initialization steps.

1. Parameterization of $\hat{o}(x)$ and $\hat{h}_m(x)$

The object estimate $\hat{o}(x)$ is always parameterized pixel by pixel in the image domain so that we do not bias the solution toward preconceived perceptions of what it should be. On the other hand, we have implemented two choices for parameterizing $h_m(x)$. The first parameterization is by its image-domain pixel values as for the object estimate. This parameterization is the most general one but can make finding global minima more difficult and produces restorations with higher variances than achievable with more restrictive parameterizations (when the more restrictive parameterizations are valid). The second parameterization, that can be used with great success as long as the spectral bandwidth of the measurements isn't too large, is in terms of the phase values of the field in the pupil of the imaging

8

system. The PSF estimates are generated from the pupil phase values with the use of the narrow-band Fourier optics model [35]. We choose to use a Zernike-based modal parameterization [30], but others have used a pixel-based parameterization [21].

2. Positivity and support

Two important pieces of additional knowledge that can be included as constraints in the estimation process are the finite support of $o(x)$, if applicable, and the fact that the intensities of $o(x)$ are positive. The support constraint is implemented simply by including in the minimization process only those pixel locations of $\hat{o}(x)$ that are in the known support. We enforce positivity by replacing $\hat{o}(x)$ in Eq.(2) with $\beta^2(x)$ and carrying out the minimization over $\beta(x)$. This parameterization of $\hat{o}(x)$ permits the minimization of Eq.(2) using unconstrained routines.

3. Regularization approaches

Two versions of Eq.(2) are implemented in PCID corresponding to the two methods of regularization. The first version sets $\gamma = 0$, so that there is not a separate regularization term. The regularization is implemented in the creation of $\hat{i}_m(x)$, which is calculated as

$$\hat{i}_m(x) = \hat{h}_m(x) * h_r(x) * \hat{o}(x),\tag{3}$$

where $h_r(x)$ is a regularization filter. A benefit of carrying out regularization as shown in Eq.(3) is that well-understood linear filters can be implemented. The potential downside of this manner of regularization is that the problem is still ill-conditioned in the sense that the global minimum of Eq.(2) is unaffected by spatial frequency values of $\hat{o}(x)$ not contained in the Fourier-domain support of $h_m(x)$. We have discovered that the inclusion of $h_r(x)$ significantly slows the amplification of these spatial frequency values of $\hat{o}(x)$, so that regularization is easily carried out by restricting the number of iterations the

minimization procedure is allowed to execute. We note that the final estimate of $o(\mathbf{x})$ is $\hat{h}_r(\mathbf{x})\hat{o}(\mathbf{x})$, not $\hat{o}(\mathbf{x})$. A second benefit of this type of regularization that follows from this fact is that, when enforcing positivity as described above, the resulting unregularized estimate satisfies the positivity constraint but the regularized version can have negatives values if the regularization filter has negative values. Although the positivity constraint can reduce noise levels in images [36] and improve convergence properties, it also limits the maximum resolution in a restored object to less than is possible for a restoration that is permitted to have negative values [37]. This regularization approach maximizes the benefits of the positivity constraint while removing its resolution limitations. The image restorations produced with this approach have excellent quality, as will be demonstrated in Section 5.

The second version of Eq.(2) that is implemented in PCID uses standard Tikhonov regularization [38]; thus, $g()$ is given by

$$ g\left[\hat{o}(\mathbf{x}),\hat{h}_1(\mathbf{x}),...,\hat{h}_M(\mathbf{x})\right] = \sum_{n=1}^{N^2} \hat{o}^2(\mathbf{x}_n). \tag{4} $$

In addition, because regularization is implemented by this penalty term, the estimated measurement is given by

$$ \hat{i}_m(\mathbf{x}) = \hat{h}_m(\mathbf{x}) * \hat{o}(\mathbf{x}). \tag{5} $$

A key benefit of using Tikhonov (or other penalty-term-based regularization schemes) is that the estimate $\hat{o}(\mathbf{x})$ obtained at convergence is the regularized version, so it is unnecessary to regularize by iteration. This property is useful when seeking to automate the algorithm execution. Our primary reason for implementing the Tikhonov penalty term, as compared to other penalty terms, is to make calculating sample variances of $\hat{o}(\mathbf{x})$ to compare with their associated CRBs computationally efficient.

Regardless of which regularization method is used, the regularization parameter must be calculated to select the amount of regularization. Many approaches to this calculation have been published (see Ref. [39] and articles referencing it). In the case of MFBD applied to imaging through atmospheric turbulence, a straightforward calculation is possible as is now described. For the linear-filter-based regularization method, we choose its cutoff frequency (if it is bandlimited) or its half-width parameter (if it is not bandlimited) in terms of where the signal-to-noise ratio (SNR) of the azimuthally-averaged image energy spectrum estimate is one. For linear image restoration using speckle imaging techniques, it has been shown [40] that the SNR of the object energy spectrum estimate is equal to the SNR of the image energy spectrum estimate. Because we have multiple measurement frames in MFBD, the image energy spectrum estimate can be calculated from the data using sample statistics.

For the second regularization approach, the image energy spectrum SNR is also calculated, but choosing and implementing the regularization parameter is more challenging. In general, Tikhonov regularization produces space-variant regularization, so a Fourier-domain filter representation of its regularization is not valid. Even in the case when Tikhonov regularization is space-invariant (white Gaussian noise), knowledge of the energy spectra of the blurring PSFs is required. We currently calculate the regularization parameter to provide the desired ratio of the data-matching term to the regularization term. In the future, we plan to calculate the regularization parameter by generating an analytical approximation to the average PSF energy spectrum [40] and approximating the white Gaussian noise parameter in terms of the measurement noise energy spectrum [9]. We then will use these values in the white Gaussian noise Tikhonov filter and choose its regularization parameter to produce the desired half-width dimension. To calculate the analytical form of the average PSF energy spectrum, the Fried parameter r_0 must be estimated from the blurred imagery [41] or from ancillary measurements.

The automatically-calculated regularization parameter provides a good starting point for the image restoration process and is especially useful for algorithm automation. The optimum regularization parameter is a function of how the restored imagery will be used; for this reason, we include the option of allowing the user to select the regularization parameter manually.

4. Minimization procedure

We use a conjugate gradient routine to minimize the cost function corresponding to the regularization approach that is implemented and all the additional knowledge that is applied. We provide to the minimization routine analytical expressions for the derivatives. The routine that we use is based on the Numerical Recipes conjugate gradient routine [42], but is significantly modified. We have removed the common block structure to make the routine less memory intensive. We have also parallelized the routine as will be discussed in Section 3. We have implemented four modifications to help speed up the minimization process. First, we choose the initial step sizes and directions for the search step in the routine based upon the MACOPT approach [43]. Second, we normalize the PSF and object derivatives to balance the impact on the cost of mismatches between the true and estimated object and PSF estimates. Third, we use half-arrays for the FFT calculations necessary to generate $\hat{i}_m(x)$ with the use of either Eq.(3) or Eq.(5). Finally, we use the FFTW software package. In the serial version of PCID, we use the 2-D version of FFTW, but in the parallelized version of PCID we use only the 1-D version since the latest official version of FFTW (FFTW 3.1.2) does not contain a distributed memory parallel implementation. An alpha version of FFTW 3.2 is available that does have a distributed-memory implementation, and we plan to evaluate its performance compared to our current implementation when it becomes the official version.

5. De-aliasing-based superresolution

It is not unusual to have measurement data that is aliased due to size of the object being imaged relative to the number of pixels in the observation camera. It is well known that de-aliasing can be accomplished if multiple sub-pixel-shifted images of the same scene are available [44]; however, the key issue in the accuracy of the de-aliased Fourier spectrum is the accuracy to which the multiple images are co-registered [45]. For an image that is aliased by a factor of X, X^2 separate and linearly-independent images must be

measured. When imaging through the atmosphere using short-exposure images, hundreds or thousands of images can be collected, so this requirement is easily satisfied. We carry out the de-aliasing procedure as part of the MLE estimation and jointly estimate the measured image shifts along with the de-aliased object and PSF restorations. The only option that needs to be selected is to embed the measured data arrays in larger arrays that are at least Nyquist sampled.

6. Two uniqueness issues

As explained in [4], there are two requirements for the support-constrained blind deconvolution problem to generate unique solutions. The first is that the sum of either the object or each of the PSF intensity values must be held constant. The reason for this requirement is that an estimated measurement remains unchanged if the object estimate is multiplied by a constant and the PSF estimate is divided by that same constant. For pixel-based blind deconvolution, it is necessary then to formulate the estimate of the mth measurement, for all m, in a way that keeps the total PSF intensity fixed. We accomplish this by replacing $\hat{h}_m(x)$ in Eqs. (3) and (5) with $\hat{h}_m(x)/\sum \hat{h}_m(x)$. This forces the sum of the object intensity values to remain fixed and thus the sum of the intensity values for each PSF. For the Zernike-based PSF estimate version of PCID, the sum of the PSF intensity values is always equal to one because of the standard mathematical model that relates the Zernike coefficients and polynomials to the corresponding PSF [35].

The second requirement for support-constrained blind deconvolution to be unique is that the locations of the object and the PSFs must be fixed. The reason for this requirement is that an estimated measurement remains unchanged if the object estimate is shifted a certain amount in a given direction and the PSF estimate is shifted the same amount in the opposite direction. We note that the structures of both the object and the PSF estimates must remain unchanged when shifted for the estimated measurement to remain unchanged. Since the PSFs for realistic imaging systems have infinite support, any shift in a PSF estimate will change its structure because of wraparound in the array containing the PSF estimate. This

13

removes, theoretically, the insensitivity of the estimated measurement to shifts in the PSF estimate. For real estimation problems, the intensity values in a PSF estimate approach zero at the edges of the array, so the blind deconvolution problem is ill-conditioned if the object and PSF locations are not fixed. We remove this ill-conditioned issue by applying support constraints to both the object and each of the PSFs. To minimize the deleterious effects of applying a support constraint to an infinite-support PSF, we choose the size of the support to encompass at least 99% of the estimated PSF intensity

7. Initialization procedures

The measurement data needs to be preprocessed to remove atmospheric background, camera dark current biases, and optical system flat field effects prior to be used in the PCID algorithm. In addition, the measurement data must be scaled so that the intensity values correspond to photon counts. If Zernike-based PSF estimation is used, the initial estimates of the Zernike parameters are set to zero, except for tip and tilt that are estimated from the measurement data. For pixel-based PSF estimation, the default initialization is an Airy pattern corresponding to the telescope used for the observations, but user-provided initial PSF estimates can be used if desired.

An initial estimate of the object is generated by co-adding centered versions of all the measurement frames. The object support region, if it is to be used as a constraint, is generated from the initial object estimate by thresholding the initial object estimate by an amount calculated from the camera read noise and the calculated r_0. The read noise is used to make sure that the object support region is not extended into the noise-only parts of the measurement frame, and the r_0 parameter is used to estimate the atmospheric blurring in order to shrink the object support to more closely match the true object support.

An important feature of the PCID algorithm is that it can use clipped frames of data that are brought about by either image jitter or by the fact that the object is too large for the imaging system field of view. As long as all parts of the object are visible when considering all the measurement frames, the PCID algorithm can restore a single unclipped estimate since it includes clipping effects in the forward model. Clipping is a problem, however, when generating either an initial estimate of the object or the

object support. For this reason, we have implemented an automatic way to detect and thus to exclude clipped measurement frames. We use the fact that the energy spectra of clipped images have excessively-high values along the frequency axis corresponding to the clipped direction to automatically identify clipped images.

3. PCID PARALLELIZATION

We describe the parallel architecture of the PCID algorithm in this section and discuss the features of the algorithm that ultimately limit its scalability across multiple processors. We then demonstrate its scalability using a Cray XD1 parallel computer that is representative of available commodity clusters, where commodity clusters are parallel computers that use non-specialized hardware. In this context, the term "scalability" refers to the ability of the PCID algorithm to use increasing numbers of processors efficiently to reduce the wall-clock execution time [46]. For example, a program can be said to scale perfectly if the wall-clock execution time decreases by a factor of two when the number of processors used increases by a factor of two. The PCID algorithm is executed on top of a message-passing interface (MPI) environment called MPICH [47].

A. PCID parallel architecture

The PCID algorithm is structured as a multiple-instruction multiple-data (MIMD) program in a master-worker architecture. The MIMD structure allows commands in a single program to be identified as belonging to the master process or a worker process and only executed by the correct process. The master process controls the overall program execution and the worker processes carry out all the computationally-intensive calculations. A top-level flowchart of the PCID parallel architecture is shown in Fig. 1 where, for clarity, only one worker process is displayed. We next describe first the master process and then the worker process.

The master process only uses one processor regardless of the number of processors available since its computational load is much lighter than for the worker processes. The master process is

responsible for the overall algorithm execution. This includes: (1) accepting all the inputs to the PCID algorithm, (2) reading in the measurement data from a file, (3) calculating all the algorithm initialization values, (4) parceling out the measurement data, the processing parameters, and the initial estimates of the object and the PSFs to the worker processes, (5) receiving back from the worker processes the information necessary update the value of the cost function and to select the next conjugate gradient direction, (6) passing back to the worker processes the updates to the estimates of the object and PSFs, and (7) determining when to halt the overall program execution either because the maximum number of iterations has been reached or because the relative change in the cost function is less than the minimum value selected. When Zernike-based PSF estimation is selected, the master process passes to each worker process the information that it needs to calculate that portion of the Zernike polynomials it will use in the estimation process.

Each worker process calculates its part of the cost function and the gradient of the cost function with respect to the current object estimate and the current PSF estimates for its measurement frames. In addition, each worker process calculates its portion of the quantities needed to determine the next conjugate direction. The calculations carried out by a given worker process are independent of all the other worker processes. This means that no communications are necessary between worker processes and thus the only communications that occur are between the master process and each of the worker processes. When each worker process uses only one processor, the communication between processors is minimal after the initialization communications and thus the limiting factors in PCID execution speed are the number and speed of the processors. For this reason, the benefit of this level of parallelization is that it scales closely with the number of processors used to run PCID. As a result, increasing the number of processors by a factor of K decreases the execution time by a factor of $1/K$ (neglecting the initialization execution time). The downside of this level of parallelization is that the maximum amount of parallelization possible is equal to the number of measurement frames. For further increases in execution speed, each worker process must use multiple processors. Issues associated with this increased level of parallelization are discussed next.
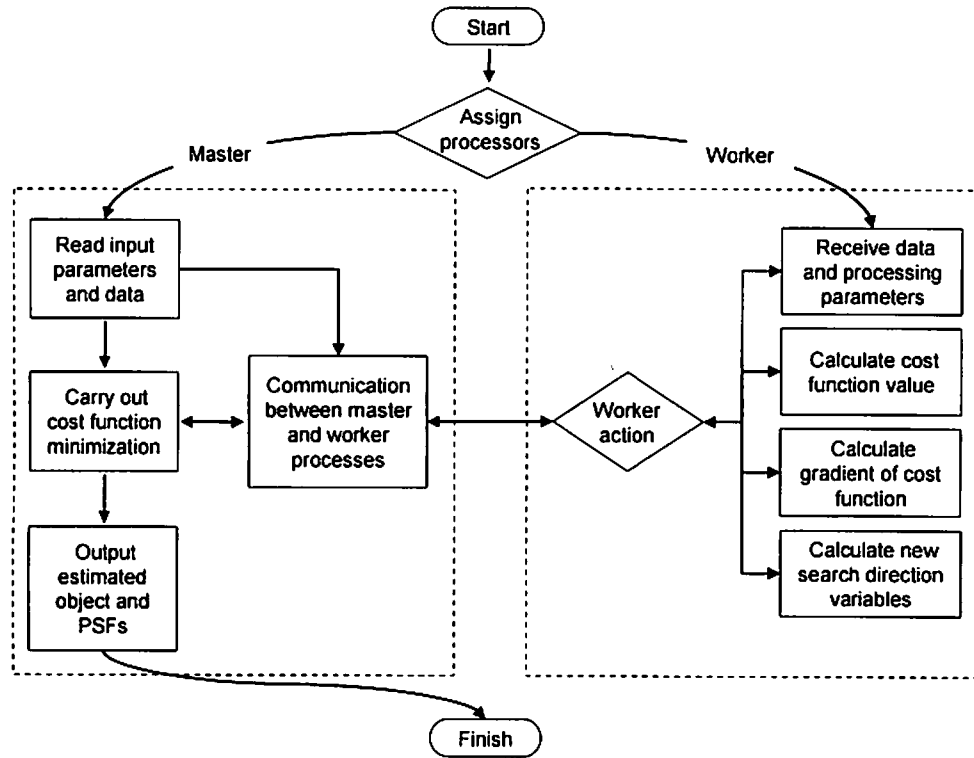
Fig. 1. PCID parallel architecture

Each worker process is required to carry out convolution operations in the course of calculating the cost function and cost function gradient. Convolution operations are carried out most efficiently in the Fourier domain. For this reason, the 2-D FFT plays a dominant role in the total execution time; in fact, profiling results show that over half of the execution time of the PCID algorithm is spent carrying out 2-D FFTs. For this reason, it is essential to maximize the speed of the 2-D FFT operation. The general 2-D FFT of an array is carried out by first replacing its rows with their 1-D FFTs, and then

replacing the columns of this array with their 1-D FFTs. When only one processor is used per worker process, all the data needed for the 2-D FFTs are contained in local memory and thus the execution speed of the 2-D FFT is driven by the speed of serially carrying out the 1-D FFTs. When multiple processors are used per worker process, the 1-D FFTs can be carried out in parallel by all the worker processors. This can increase the execution speed of the 2-D FFT; however, data transfer is necessary between the worker processors during the 2-D FFT operation which mitigates the speed improvements, even when high-speed interconnects such as Infiniband [48] and Myrinet [49] are employed. For this reason, we have sought to minimize the communications between the processors in a worker process to carry out the 2-D FFT. Because all fast 1-D FFTs require the data to be in local memory, the master process always spreads the data across worker processors by row so that the first set of 1-D FFTs can be executed with the rows currently in memory. To maximize the speed of execution, each worker process must have the same number of rows. After the 1-D row FFTs have been calculated, the rows of the array must be redistributed among the worker processors so that each column of the array resides in the local memory of a worker processor. To explain how we do this, suppose that there are two processors in a worker process as shown in Fig. 2. For an N by N array (N even), one processor is sent the first N/2 rows and the second is sent the last N/2 rows. After 1-D FFTs are taken of each row, the resulting rows are split into two N/2 by N/2 blocks that are separately transposed. All of these operations take place in local memory, so no communication is necessary between the worker processors. Next, the off-diagonal blocks are swapped between the two processors – this is the only step that requires communication between the two worker processors. Once the transfer is complete, each worker processor has N/2 columns of the array and can complete the 2-D FFT in local memory. For the general case of W worker processors per worker process, N must be evenly divisible by W. Because rows and columns are swapped, communications only occur between pairs of worker processors, permitting the use of point-to-point MPI communications.
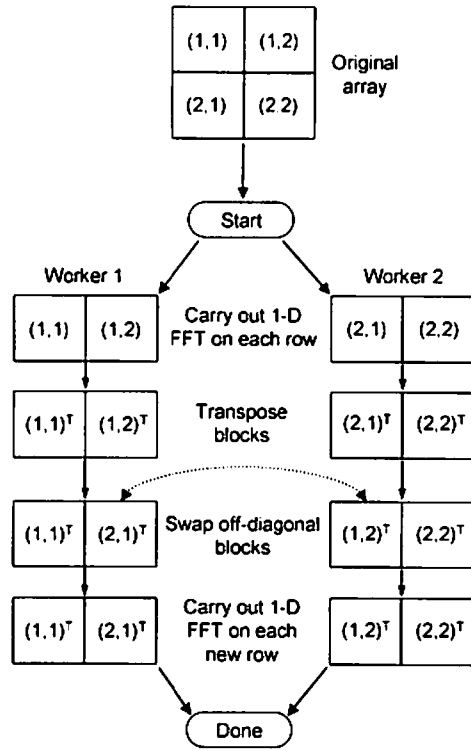
Fig. 2. PCID parallel 2-D FFT flowchart

Since each processor must have, at a minimum, an entire row of the array to be Fourier transformed using fast 2-D FFT routines, the maximum number of processors that can be used per measurement frame is N. When M measurement frames are used for a single object estimate, the maximum number of processors that PCID can use is MN. One example of how many processors that can be used by PCID is for the case of arrays that are 1024 by 1024, and when 400 measurement frames are used, as is the situation for one of the example restorations in Section 5. Theoretically, 409,600 processors can be used in the restoration process; however, internode communication time dominates the algorithm execution time for far fewer processors.

B. PCID parallel scalability demonstration

We used a Cray XD1 supercomputer [33] located at the Maui High-Performance Computing Center [50] to explore the scalability of PCID as a function of the number of worker processors used in a worker process. The Cray XD1 is an SLES 9.0 Linux-based cluster with 144 nodes, each with two 2.4 GHz AMD64 Opteron processors and 8 GBytes of RAM. Each processor has a theoretical peak speed of ~4.8 GFLOPS. The nodes are connected via a Fat Tree RapidArray Interconnect switch that has a bandwidth of 4 GBytes/second with a MPI latency of 1.7 μs. The PCID algorithm uses the MPI standard to communicate between nodes. For maximum speed, it uses vendor-optimized versions on the platforms that it runs on when available (as is the case for the Cray XD1), but is fully compatible with both the MPICH1 and MPICH2 freely-available implementations of MPI [47]. Although the Cray XD1 nodes have two processors each, we chose to use only one processor per node for the scalability tests to ensure that all interprocessor communications use the internode communication channels, not on-board communication channels. This is the severest test of the scalability of a parallel program.

Execution times for PCID as a function of the number of processors included in a worker process for measurement frames that are 128 by 128 pixels in size are shown in Fig. 3a. The numbers of processors included in the worker process were limited to powers of two because the array size is a power of two. These results were generated using only one worker process and letting PCID run for 100 iterations (one iteration corresponds to a single conjugate-gradient search direction). It can be seen that, as the number of processors is increased from 1 to 32, the PCID execution times decrease, indicating that the execution times in this region are limited by the availability of processor cycles. For more than 32 processors, the execution times increase as the number of processors is increased, indicating that internode communications are limiting the execution time. To quantify the scalability of PCID across processors, we calculated a scalability factor for each number of processors greater than one using the data in Fig. 3a. For a given number of processors, say 2^n, the scalability factor is equal to the execution time for 2^{n-1} processors divided by the execution time for 2^n processors. If the PCID algorithm scaled perfectly, the scalability factor would be equal to two for any value of n. If adding processors did not

help, the factor would be one. Values of the factor less than one indicate that execution times are increasing as processors are added. The scalability factor is plotted in Fig. 3b as a function of the number of processors. It can be seen that PCID only scales perfectly when moving from one to two processors. Beyond this, the scalability factor drops to around 1.5. Finally, for more than 32 processors, the scalability factor is less than one because execution times are increasing.
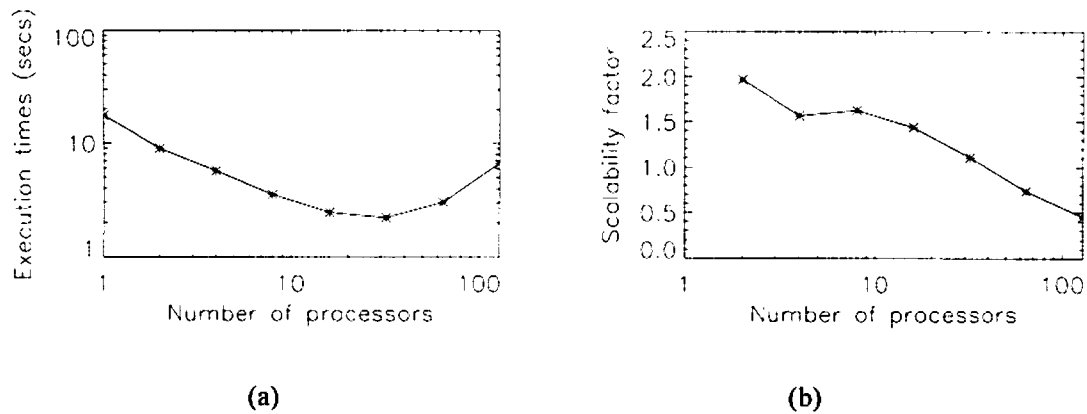

(a)

(b)

Fig. 3. (a) Execution times and (b) scalability factors for the PCID algorithm for 128 by 128 measurement frames.

We next generated results in the same way as for the results in Fig. 3 except we used measurement frames that are 512 by 512 pixels in size. The results are shown in Fig. 4. It can be seen that the scalability for larger arrays is better than for smaller arrays. This is due to the fact that, for larger arrays, each processor has more data in local memory than for smaller arrays, thus the execution time is limited by the CPU cycles available. Eventually the internode communications load will cause execution times to increase as more processors are added.
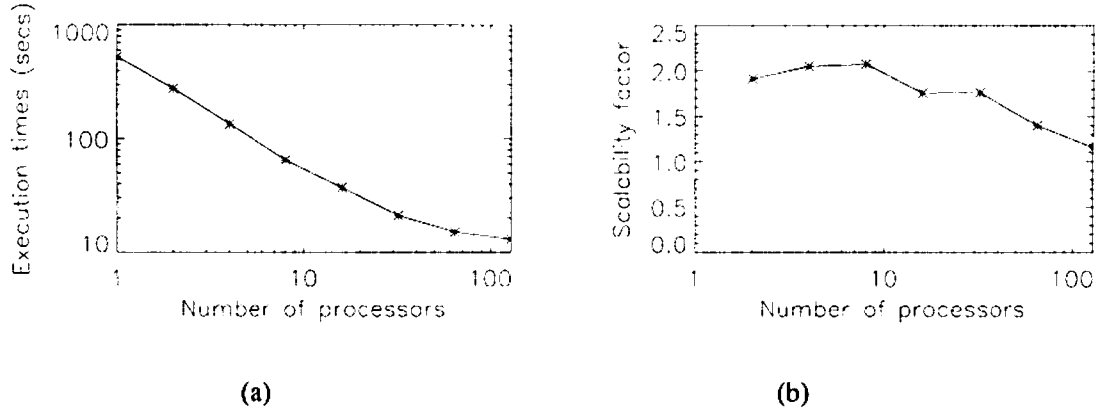
Fig. 4. (a) Execution times and (b) scalability factors for the PCID algorithm for 512 by 512 measurement frames.

As a reminder, the results shown in Figs. 3 and 4 demonstrate the scalability of adding multiple processors to a worker process for a single measurement frame. If there are multiple measurement frames to be included in the restoration process, the best parallelization approach is data parallel if possible, since the scalability is essentially perfect in this case. Only when there are more processors than measurement frames should one include multiple processors in a worker process.

4. CRB THEORY AND RESULTS

To evaluate the quality of image restorations using a given algorithm, it is useful to have algorithm-independent limits to image quality as benchmarks. There are a number of algorithm-independent metrics that can be used to evaluate image quality, many of which are subjective in nature. We choose to quantify image quality in terms of the SNRs of the image restorations in order to have an objective metric to evaluate algorithm performance. Cramér-Rao lower bound theory provides a way to calculate algorithm-independent lower bounds (CRBs) to the variances of estimates of a set of parameters (thus upper bounds to the SNRs of parameter estimates). These lower bounds can be calculated for unbiased as

well as biased estimates of the parameters. In this section we present an overview of CRB theory for unbiased and biased estimates of a set of parameters and then apply the theory to unbiased as well as regularized (a specific type of bias) estimates of the image-domain intensities of an object. In particular, we present CRB expressions for scenarios that we also can calculate meaningful sample variances using PCID. We compare these CRBs to the PCID sample variances and show that image restorations produced by PCID achieve or closely approach them.

A. CRB theory

Let Φ be a vector containing the parameters to be estimated. Any unbiased estimate of Φ, $\hat{\Phi}_u$, that is based on a set of measurements y has variances that are bounded below by the CRBs that are the diagonal elements of the inverse of the Fisher information matrix (FIM) F corresponding to Φ and y. Mathematically, we have

$$\mathrm{var}\left(\hat{\Phi}_u\right) \geq \mathrm{diag}\left(F^{-1}\right), \tag{6}$$

where $\mathrm{var}\left(\hat{\Phi}_u\right)$ is the variance of $\hat{\Phi}_u$ and $\mathrm{diag}(F^{-1})$ denotes the diagonal elements of F^{-1}. The element of F in the pth row and q^{th} column, $[F]_{pq}$, is given by

$$[F]_{pq} = E\left\{\frac{\partial \ln[f(y;\Phi)]}{\partial \Phi_p}\frac{\partial \ln[f(y;\Phi)]}{\partial \Phi_q}\right\}, \tag{7}$$

where ln is the natural logarithm, $f(y;\Phi)$ is the probability density function of the measurement y that is a function of Φ, and E{ } is the expected value of the quantity in braces.

Often, biased estimates of Φ are sought because either unbiased estimates are not possible or because a given biased estimate has more desirable properties than an unbiased estimate such as lower

mean-square error. Let $\hat{\Phi}_b$ be a biased estimate of Φ with bias b_Φ, where the subscript Φ indicates that the bias is a function of the true parameters. Then the variances of $\hat{\Phi}_b$ are bounded below as shown in the following equation:

$$\text{var}\left(\hat{\Phi}_b\right) \geq \text{diag}\left[\left(I + \nabla_\Phi b_\Phi\right)F^{-1}\left(I + \nabla_\Phi b_\Phi\right)'\right],\tag{8}$$

where $\nabla_\Phi b_\Phi$ is the gradient of b_Φ with respect to Φ and I is an identity matrix of the appropriate dimension.

To apply CRB theory to the MFBD image restoration problem, the imaging model of Eq.(1) must be rewritten in a discrete vector form. To this end, let α be a vector that contains the spatial locations of the intensity values of $i_m(x)$ as would be obtained for an image collected with a CCD array. Then let y_m, θ, and η_m be one-dimensional vectors that contain the values of $i_m(\alpha)$, $o(\alpha)$, and $n_m(\alpha)$, respectively, on the grid defined by α. In addition, let H_m be the block-circulant system matrix associated with $h_m(\alpha)$ [51]. This permits rewriting Eq.(1) as a matrix-vector equation given by

$$y_m = H_m\theta + \eta_m.\tag{9}$$

Now let all M measurements be concatenated into a single measurement vector y, where $y = [y_1,....,y_M]^T$. Then, using Eq.(9), we can write the concatenated measurement equation as

$$y = H\theta + \eta,\tag{10}$$

where $H\theta = [H_1\theta_{...}, H_N\theta]^T$, $\eta = [\eta_1, ..., \eta_M]^T$ and the PDF of η is denoted by $f_\eta(\eta)$. Because both photon and camera read noises are statistically independent across pixels and across measurement frames, $f_\eta(\eta)$ is equal to the multiplication of all the single-pixel PDFs.

Using Eqs. (6) - (8) and (10), we can derive CRB expressions for unbiased and biased estimates of restored images that can be compared to the corresponding PCID restored-image sample variances; however, prior to doing so, two issues must be addressed. The first issue is that an unbiased estimate of the underlying true object must be attainable. For a realistic imaging system, an unbiased estimate is not attainable because its PSF is not invertible. Theoretically, an unbiased estimate is attainable when utilizing a support constraint as a result of superresolution even with a non-invertible PSF; however, with any noise at all, all superresolved Fourier spectra are biased [52]. Nonetheless, we will derive unbiased image-restoration CRBs for two reasons. The first is that we can employ invertible PSFs in order to test the performance of the PCID algorithm. The second reason is that we can use a particular form of the FIM for non-blind unbiased CRBs to calculate non-blind regularized CRBs for non-invertible PSFs for regularizing filters that are admissible (as defined in the next paragraph).

The second issue is the calculation of biased CRBs. It can be seen from Eq. (8) that an analytical expression for the gradient of the bias must be available to calculate biased CRBs. Such an expression is available for regularized image restorations when the Fourier-domain support of the regularization filter is contained in the Fourier-domain support of the Fourier-transformed PSF but is generally not available otherwise. (We will refer to this type of regularization filter as an admissible regularization filter). It is not available for any type of regularization that includes estimated Fourier spectra outside the passband of the PSF, including all penalty-based regularization methods such as Tikhonov regularization. The reason that analytical expressions are not available for non-admissible regularization is because this type of regularization passes restored frequency content outside the bandwidth of the measurement system; that is, they pass superresolved frequency content. As stated in the previous paragraph, all superresolved

Fourier data are biased, and analytical expressions for the biases are not known except in limited situations.

Another situation where bias gradient expressions are not available is when inequality constraints such as positivity are used in the image restoration process. Initial steps have been taken to generate an alternate way to calculate inequality-constrained biased CRBs [36], but much more work needs to be carried out before this approach can be applied to positivity-constrained image restoration.

Because of these two issues, there are only a limited number of scenarios for which both PCID image restorations can be generated and CRBs can be calculated. A support constraint can always be included in any of these scenarios. All scenarios that include unbiased image restorations fit in this category. Almost no scenarios that include biased image restorations fit in this category because of the two ways that PCID implements regularization. The Fourier-filter-based regularization approach accomplishes the requisite regularization by limiting the number of iterations. This produces biases in the image restorations that can not be calculated analytically. The Tikhonov regularization approach allows the algorithm to iterate to convergence; however, the image restorations contain spatial frequency content from the entire Fourier domain. As a result, only scenarios that include invertible PSFs have biases that can be calculated analytically. In addition, our CRB calculations assume space-invariant regularization, so only white measurement noise can be modeled when Tikhonov regularization is employed [38].

As a result of the issues discussed above, we have been able to compare sample variances of PCID image restorations to their associated CRBS for the following scenarios: (1) non-blind unbiased image restorations for any combination of read and photon noises, (2) non-blind biased image restorations for invertible PSFs, Tikhonov regularization, and white read noise, and (3) blind unbiased image restorations for pixel-based invertible PSF estimation for any combination of read and photon noises. No blind image restorations that use Zernike-based PSF estimation are included in the scenarios because all Zernike-based PSFs are non-invertible. Any support constraints that contain the true support regions for the object and/or PSFs can be included in any of these scenarios.

Next, we present an expression for the FIM to be used in Eqs. (6) and (8) to generate the CRBs that will be compared to PCID sample variances. We present the FIM for blind deconvolution with pixel-based PSF estimation since the FIM for non-blind deconvolution is a special case of the blind FIM. In blind deconvolution, the functionally-independent parameters that generate the elements of both $\{H_m\}$ and θ are estimated, where $\{H_m\}$ denotes the set H_m, m=1,...,M. Let Φ be the vector that contains all of these parameters. For pixel-based estimation of $h_m(\alpha)$ and $o(\alpha)$ when $\sum_{n=1}^{N} h_m(\alpha_n)$ is kept fixed, only N-1 elements of each of the $h_m(\alpha)$ are functionally independent. Without loss of generality, we assume that the first N-1 elements of each of the $h_m(\alpha)$ are used to generate each corresponding H_m and thus are the elements of the vector Φ along with the N elements of θ. The FIM, F, associated with y, Φ, and η has a block structure that is given by

$$
F = \begin{bmatrix} F_{11} & \cdot & \cdot & \cdot & F_{(M+1)1} \\ \cdot & \cdot & & 0 & \\ \cdot & & \cdot & & \\ \cdot & 0 & & \cdot & \\ F_{(M+1)1} & & & & F_{(M+1)(M+1)} \end{bmatrix},
\qquad (11)
$$

where $F_{m1} = F_{1m}$ (i.e., the FIM is symmetric). In terms of this block structure, the first column, the first row, and the diagonal are the only parts of F that are non-zero. Using the results in [53], it can be shown that the element of F_{11} in the pth row and the qth column, $[F_{11}]_{pq}$, is given by

$$
[F_{11}]_{pq} = \sum_{m=1}^{M} \sum_{n=1}^{N^2} \frac{h_m(x_n - x_p)h_m(x_n - x_q)}{i_m(x_n) + \sigma_m^2(x_n)}
\qquad (12)
$$

where $\sigma_m^2(x)$ is the variance of the read noise in the mth image. Similarly, the element of F_{m1} in the pth row and the qth column, $[F_{m1}]_{pq}$, is given by

$$\left[\mathbf{F}_{m1}\right]_{pq} = \sum_{n=1}^{N^2} \frac{h_m\left(\mathbf{x}_n - \mathbf{x}_p\right)\left[o\left(\mathbf{x}_n - \mathbf{x}_q\right) - o\left(\mathbf{x}_n - \mathbf{x}_{N^2}\right)\right]}{i_m\left(\mathbf{x}_n\right) + \sigma_m^2\left(\mathbf{x}_n\right)} \, . \tag{13}$$

Finally, the element of $\mathbf{F}_{mm}$ in the pth row and the qth column, $[\mathbf{F}_{mm}]_{pq}$, $m \geq 2$, is given by

$$\left[\mathbf{F}_{mm}\right]_{pq} = \sum_{n=1}^{N^2} \frac{\left[o\left(\mathbf{x}_n - \mathbf{x}_p\right) - o\left(\mathbf{x}_n - \mathbf{x}_{N^2}\right)\right]\left[o\left(\mathbf{x}_n - \mathbf{x}_q\right) - o\left(\mathbf{x}_n - \mathbf{x}_{N^2}\right)\right]}{i_m\left(\mathbf{x}_n\right) + \sigma_m^2\left(\mathbf{x}_n\right)} \, . \tag{14}$$

Equations (12) – (14) give the elements of F without support constraints. To include support constraints, the elements of Φ are restricted to contain only the locations of $\{h_m(\alpha)\}$ and $o(\alpha)$ that are inside the support constraint regions, where $\{h_m(\alpha)\}$ denotes the set $h_m(\alpha)$, for m=1,...,M. In addition, the non-blind version of F is equal to $[\mathbf{F}_{11}]$.

The above expression for F can be used to generate the CRBs for unbiased estimates of the object and PSF intensities with the use of Eq. (6). To generate regularized CRBs, the bias gradient term, $\mathbf{I} + \nabla_\Phi \mathbf{b}_\Phi$, must be calculated for the regularizing filter that will be used. The general form of the bias gradient term is block diagonal as shown below:

$$\mathbf{I} + \nabla_\Phi \mathbf{b}_\Phi = \begin{bmatrix} \mathbf{G}_{11} & & & \\ & \cdot & \mathbf{0} & \\ & & \cdot & \\ & \mathbf{0} & & \cdot \\ & & & \mathbf{G}_{(M+1)(M+1)} \end{bmatrix}, \tag{15}$$

where the $\mathbf{G}_{11}$ block is used to generate the regularized image CRBs, and the $\mathbf{G}_{mm}$ block, $m \geq 2$, is used to generate regularized CRBs for the $(m-1)^{th}$ PSF. Our interest is in regularized image restorations, not regularized PSF restorations. For this reason, we set all blocks of the bias gradient term to zero except for

$\mathbf{G}_{11}$. For a specified regularizing PSF $h_{reg}(\mathbf{x})$, the element of $\mathbf{G}_{11}$ in the pth row and the qth column, $[\mathbf{G}_{11}]_{pq}$, is given by

$$[\mathbf{G}_{11}]_{pq} = h_{reg}(\mathbf{x}_p - \mathbf{x}_q).$$ (16)

B. Comparison of PCID sample variances and CRBs

In this section we present the results of comparing sample variances of parameter estimates generated using PCID to their associated CRBs to show how closely the PCID sample variances approach their theoretical limits. For unbiased image restorations, the parameters are the true object intensity values for each pixel in the object support. For regularized image restorations, the parameters are the regularized object intensity values in the object support. Other sets of parameters could have been chosen; for example, the comparison could have been made in terms of the power spectra of the restored imagery. We chose to carry out the comparison in terms of intensity values in the restored imagery because this is a useful and convenient metric. Since functional transformations of parameters have corresponding functional transformations of their CRBs [25], the selection of a specific parameter set will change the absolute values of the sample variances and the CRBs, but not the results of the comparison. The comparisons in this section are carried out for the three scenarios described in Section 4.A.: (1) non-blind unbiased image restorations corrupted by photon and/or read noises, (2) non-blind biased image restorations for invertible PSFs, Tikhonov regularization, and white read noise, and (3) blind unbiased image restorations corrupted by photon and/or read noises for pixel-based invertible PSF estimation. We show that the PCID image restorations have sample variances that either achieve or closely approach their associated CRBs.

To carry out the sample variance/CRB comparison for a given imaging and image restoration scenario, we first generated 50 measurement data sets that differed only in their realizations of the read and photon noise values. Most of the measurement data sets consisted of one frame of data, but a few of

them contained ten frames of data. We then used PCID to generate 50 restored images, one for each measurement data set.. Next we calculated sample variances and biases, pixel by pixel, using the 50 restorations. For unbiased restorations, we calculated the biases in terms of the true object being imaged. For regularization-based biased restorations, we calculated the biases in terms of the regularized version of the true object. Because of the way that we calculated the biases, they should be negligible relative to the variances. We used this fact to ensure that the PCID image restorations correspond to the true global-minimum solutions. Once this was verified, we calculated the ratio of the sum of the sample variances to the sum of the CRBs. We refer to this ratio as the PCID/CRB ratio. Numbers close to one show that the PCID image restorations are achieving the CRBs. Because of the large number of comparisons that underlie the results in this section, only the minimum, maximum, and mean values of this ratio are presented for each of the three scenarios described above.

The two true objects used in the comparison study, shown in Fig. 5, are a computer-simulated satellite model called OCNR and a two-disk object called two-circ. The satellite object was chosen because of its high dynamic range and significant level of detail. The two-disk object was chosen to be a counterpoint to the satellite model by having only two intensity levels (one in each disk) and a simple structure.
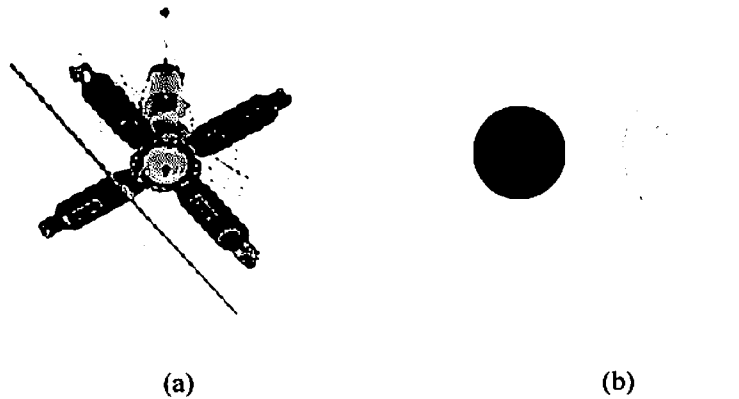
(a) (b)

Fig. 5. The two objects used in the PCID sample variance and CRB comparison: (a) OCNR, (b) two-circ.

Two PSFs were used for the comparison, both invertible as required for the comparison. The first PSF, called tripsf, is the inverse Fourier transform of the following function H(f):

$$H(f) = (1 - 0.97 f_1 / f_c)(1 - 0.97 f_2 / f_c), \tag{17}$$

where $f = (f_1, f_2)$ and f_c is the maximum spatial frequency in one dimension in the arrays containing the true objects. Because H(f) has no zeros, it produces an invertible PSF. The second PSF, called atmpsf, was created as follows. First, we generated a computer-simulated image of an unresolved star as seen through turbulence and an imaging system characterized by the ratio $D/r_0 = 8$, where D is the diameter of the imaging system [9]. This image is a PSF, but it is non-invertible because the imaging system was modeled realistically. We converted it to being invertible by multiplying it by a circular support region with a radius of 10 pixels that contained 99% of the true PSF's energy. For some of the image restorations, we used ten measurement frames with this style of turbulence PSF. In this situation, ten separate PSFs were generated, all with the same D/r_0 ratio, but with separate atmospheric turbulence realizations. The PSF corresponding to Eq. (17) and one of the turbulence PSFs are shown in Fig. 6.

The three object support constraint options available for use by PCID were the true object support, a larger version of the true object support created by blurring the true support with a 2 by 2 blurring kernel (blur2 support), and the smallest centered circular support region that contained the true object (circle support). When carrying out blind deconvolution using PCID, only the atmpsf PSFs were used. For these PSFs, the three PSF support constraint options were the true support, a centered circle of radius 20 pixels, and a centered circle of radius 30 pixels.

For scenario 1, the two noise options were read noise only and photon plus read noise. The latter noise model was dominated by the photon noise component in total, but the read noise component was included to eliminate the possibility of dividing by zero in pixels where the photon noise component was

zero (see Eq. (12)). For scenario 2, only read noise was modeled for the reasons given previously. For scenario 3, both noise options were used.

The regularized results in scenario (2) used a value of the Tikhonov weighting parameter $\gamma = 0.1$. The Fourier-domain filter corresponding to white Gaussian noise with the measurement PSF and this weighting parameter was generated using the expression in [38]. The use of this expression required that only one measurement frame be included in each image restoration.

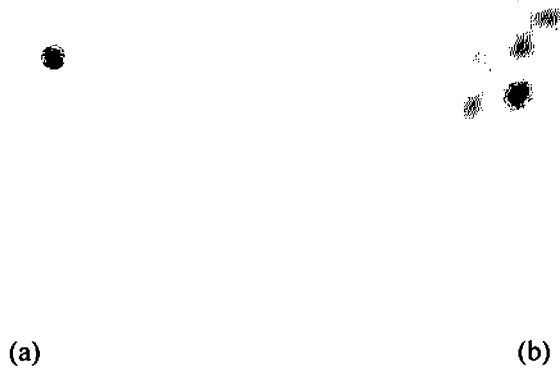(a)   (b)

Fig. 6. The two PSFs used in the PCID sample variance and CRB comparison study: (a) tripsf, (b) one realization of atmpsf. These images are magnified in size by a factor of 8 as compared to the images shown in Fig. 5 in order to better display their detail.

The minimum, maximum, and mean values of the PCID/CRB ratios for all three scenarios are shown in Table 1. Theoretically, the values of the PCID/CRB ratios should be no less than one for any image restoration; however, due to statistical variations in the sample variances, some of the ratios are

slightly less than one. The scenario 1 results were generated from 24 separate combinations of objects, object supports, PSFs, and types of noise. All the results used one measurement frame. The PCID/CRB ratios most closely clustered around one for scenario 1, clearly showing that PCID achieved the CRBs for all the image restorations. The scenario 2 results were generated using the read-noise-only subset of 12 image restorations used for the scenario 1 results. The spread of the PCID/CRB ratios is greater than for scenario 1, but the mean is still essentially one. The scenario 3 results were generated using only the atmpsfs, ten measurement frames per image restoration, and the true and blur2 object supports. Because scenario 3 is the blind-deconvolution scenario, the PSF support constraints described above were also used. Because either the object support constraint or the PSF support constraint must be the true support to keep the blind deconvolution problem unique, image restorations had at most one support constraint larger than the true support. Sixteen combinations of these options were used to generate the scenario 3 results in Table 1. The PCID sample variances are higher than for scenarios 1 and 2, but they still are only 8% higher than the CRBs

Table 1. Mean, minimum, and maximum values of the PCID/CRB ratios for the scenarios described in the text.

Scenario	PCID/CRB Ratios		
	Mean	Minimum	Maximum
1	0.9854	0.9402	1.0326
2	1.0135	0.9499	1.0732
3	1.0804	0.9791	1.205

To achieve the CRBs, the global minimum of the cost function must be found. As the noise levels increase, and as the atmospheric blurring increases, it becomes more difficult to find the global

minimum and thus to achieve the CRBs. We are working on ways to increase the robustness of PCID in these stressing scenarios.

5. IMAGE RESTORATION EXAMPLES

In this section we show image restorations of two space objects to demonstrate the quality of restorations possible with PCID. Both objects were imaged using telescopes that are part of the Maui Space Surveillance System (MSSS) [54]. The two objects, the Space Shuttle and the Hubble Space Telescope (HST), were chosen for the image restorations because they are familiar to a wide audience and because good detail can be seen in the restorations. The use of multiple images of the same object increases both the restored image quality and the probability that the PCID algorithm will find the global minimum of the cost function shown in Eq. (2); however, the unblurred object has to be the same in all the measurement frames. For space objects orbiting the earth at low altitudes such as the Space Shuttle and the HST, typical observation times for which the aspect ratio of the object remains unchanged is just a few seconds. The restorations presented in this section use all the measurement frames available within the allowed observation times.

5.1 Space Shuttle image restorations

The measurement data for the Space Shuttle restorations were collected with a camera mounted on the MSSS 1.6 meter telescope without the use of adaptive optics compensation. Representative measurement frames are shown in Fig. 7. The camera has a silicon-based CCD detector that is 128 by 128 pixels in size, four readout channels to increase the frame rate, and a read noise of 4 electrons. The spectral filter through which the light was collected is 120 nm in width with a mean wavelength of 940 nm. The exposure time for each measurement frame was 0.9 msec, ensuring that the atmospheric turbulence blurring was essentially fixed in each frame. The camera frame rate was 250 Hz, so several hundred measurement frames were able to be included in a single image restoration. Because of the large angular extent of the Space Shuttle, the telescope's field of view (FOV) was set to 120 microradians. At the mean

34

observation wavelength of 940 nm, the camera's sampling rate was only 30% of the Nyquist rate, resulting in data that was aliased by a factor of almost four. In addition, the Space Shuttle was slightly larger than this FOV even when perfectly centered. Due to residual atmosphere and telescope jitter, many of the raw images are clipped a significant amount. It will be described later in this section how these negative effects are accounted for in the image restoration process.

Another issue associated with the angular size of the Space Shuttle is that it is many times larger than the isoplanatic angle (the angular extent over which atmospheric turbulence can be modeled as space invariant). The PCID forward model assumes that the image can be modeled as a convolution of the atmospheric blurring and the unblurred object, an assumption not satisfied by the data. As will be shown, though, the image restorations are excellent and contain near-diffraction-limited detail. This indicates that the PCID algorithm is robust with respect to violations of the isoplanatic assumption governing the forward model of the measurement.

To account for the clipping artifacts in the measurement frames, the 128 by 128 measurement frame arrays were embedded in the center of 256 by 256 arrays with zero values in the extra pixels. The embedded measurement frames now have a FOV of 240 microradians. The embedding process, of course, does not remove the clipping in the measurements. What it does do is permit the restoration of an unclipped version of the Space Shuttle in a 256 by 256 array since the forward model in PCID relating the object estimate and each PSF estimate to the measured data incorporates the120 microradian FOV and the size of the Space Shuttle easily fits in a 256 by 256 array.

To make possible unaliased image restorations, each embedded measurement frame was upsampled from 256 by 256 pixels to 1024 by 1024 pixels by representing a single pixel in the 256 by 256 array with a 4 by 4 block in the 1024 by 1024 array, where each value in the 4 by 4 block is the value in the single pixel. Each upsampled measurement frame was then spatially smoothed with a 4 by 4 smoothing kernel. For this amount of aliasing, a minimum of 16 linearly-independent data frames are needed to accomplish the de-aliasing, a limit exceeded by more than a factor of ten for the restorations in this section.
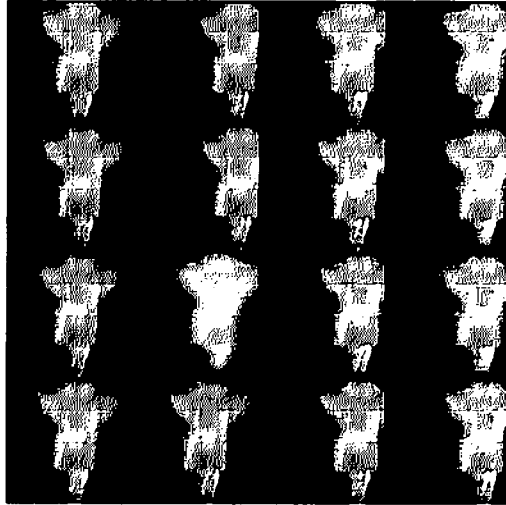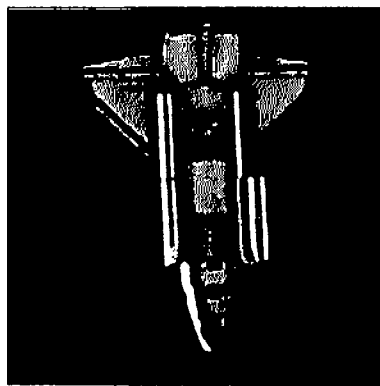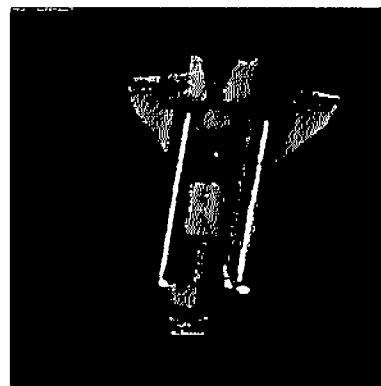
Fig. 7. Representative measurement frames of the Space Shuttle.

(a)   (b)

Fig. 8. Two image restorations of the Space Shuttle produced by the PCID algorithm.

The PCID algorithm was configured to estimate the PSFs in each measurement frame using the Zernike-polynomial parameterization option. The first one hundred Zernike polynomials were used in the estimation process. All of the Zernike coefficients were initialized to zero except for the tip and tilt values that were initialized for each PSF to the tip and tilt values in the corresponding measurement frame. A support constraint for the object restoration was generated automatically as described in Section 2 and used in the restoration process as was a positivity constraint.

Two restored images of the Space Shuttle from two separate parts of the pass are shown in Fig. 8. The restoration in Fig. 8a was achieved with the use of 400 measurement frames. Notice the high quality of the restoration as demonstrated by the detail seen in the cargo bay and on the nose. In particular, the two round windows right next to the cargo bay are resolved to near the diffraction limit of the optical system. The restored image shows all parts of the Space Shuttle with no clipping because every part of the Space Shuttle is visible in at least one measurement frame used in the restoration process. The restored image of the shuttle in Fig. 8b was created using 300 measurement frames. It is of lesser quality because of the fewer number of frames. Notice that the tip of the nose is clipped in the restoration as a result of it being clipped in all the measurement frames used for the restoration. Even though the nose is clipped, the rest of the restored image is unaffected by the clipping artifact.

5.2 HST image restorations

The measurement data for the HST restorations were collected using the MSSS AEOS 3.6 meter telescope with its visible-spectrum adaptive optics system [55]. Two representative measurement data frames are shown in Fig. 9. The adaptive optics system has a 941 actuator deformable mirror and a Shack-Hartmann-based wavefront sensor. The camera used to record the data has a silicon-based CCD detector that is 512 by 512 pixels in size, only one readout channel, and a read noise of 12 electrons. The spectral filter through which the light was collected is 360 nm in width with a mean wavelength of 880

nm. The exposure time for each measurement frame was 98 msec. The exposure time was much greater than the atmospheric correlation time; however, because of the adaptive optics correction, the effective system correlation time was much greater than the atmospheric correlation time. The camera frame rate was 2.4 Hz, so that only six measurement frames were included in a single image restoration. The FOV of the telescope was set to 51 microradians, large enough to contain the HST and small enough to produce Nyquist-sampled data. In addition, the degree of anisoplanicity was much less than for the Space Shuttle image restorations.

As for the Space Shuttle image restorations, the PCID algorithm was configured to estimate the PSFs in each image using the Zernike-polynomial parameterization option. The first one hundred Zernike polynomials were used in the estimation process. All of the Zernike coefficients were initialized to zero except for the tip and tilt values that were initialized for each PSF to the tip and tilt values in the corresponding measurement frame. A support constraint for the object restoration was generated automatically as described in Section 2 and used in the restoration process as was a positivity constraint.
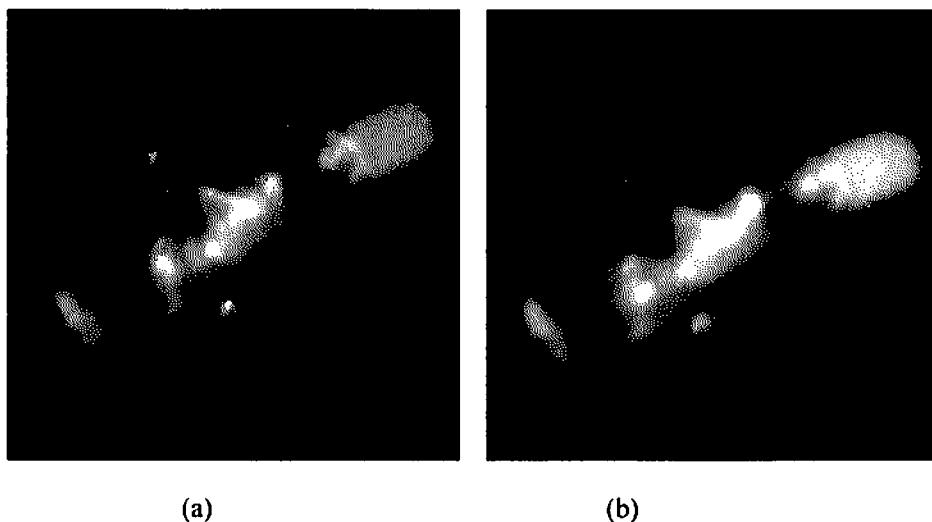

(a)   (b)

Fig. 9. Two representative measurement frames of the HST.
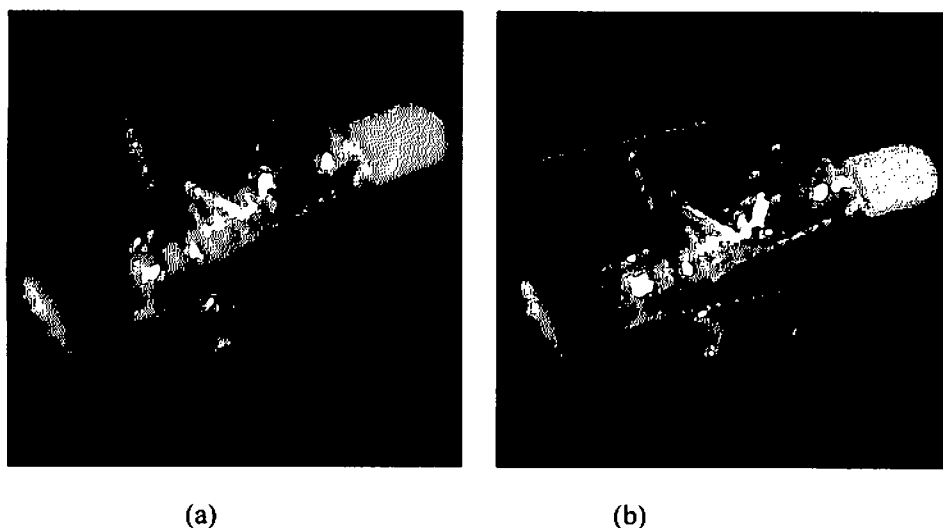
(a)   (b)

Fig. 10. Two PCID restorations of the HST.

Two HST restorations are shown in Fig. 10. There is a noticeable increase in image quality in these restorations when compared to the representative measurement frames; however, the amount of improvement is less than seen in the Space Shuttle restorations, for two reasons. First, the HST measurement frames have much higher quality than the Space Shuttle measurement frames even before processing. As a result, there is less room for improvement before reaching the diffraction limit of the telescope. Second, far fewer frames were used for the HST restorations than for the Space Shuttle restorations.

6. CONCLUSIONS AND FUTURE WORK

We have presented our blind deconvolution algorithm, called PCID, described its features, and characterized its performance. Although many blind deconvolution algorithms have been developed and published in the past, several features of our algorithm extend the state-of-the-art in blind deconvolution.

First, we have parallelized the algorithm so that it can run on commodity clusters – specialized hardware is not required. We demonstrated the speed improvements using a Cray XD-1 cluster and showed how PCID's scalability properties depend on the size of the data arrays. We have also developed Cramér-Rao lower bound expressions for blind deconvolution and carried out an extensive comparison study that showed that PCID closely approaches these theoretical limits. Finally, we presented image restorations of the Space Shuttle and the Hubble Space Telescope from data collected by two separate telescopes, with and without adaptive optics, and showed that significant improvements in image quality were obtained.

Future work includes implementing the complete MACOPT [43] conjugate gradient minimizer which should noticeably increase PCID's execution speed. For further speed improvements, we plan to implement the distributed-memory version of FFTW when it becomes available should it prove to be faster than our current 2-D FFT. We will further increase the automatic parameter calculation properties of PCID to make it more user friendly. We are also in the process of writing a web-based GUI that will be a front end to PCID. In addition, this GUI will be used for data I/O, job submission, and data visualization.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. Kundur and D. Hatzinakos, "Blind image deconvolution," IEEE Signal Process. Mag. 13(3), 43-64 (1996).

[2] D. Kundur and D. Hatzinakos, "Blind image deconvolution revisited," IEEE Signal Process. Mag. 13(6), 61-63 (1996).

[3] T. G. Stockham, Jr., T. M. Cannon, and R. B. Ingebretsen, "Blind deconvolution through digital signal processing," Proc. IEEE **63**, 678-692 (1975).

[4] R. G. Lane and R. H. T. Bates, "Automatic multidimensional deconvolution," J. Opt. Soc. Am. A **4**, 180-188 (1987).

[5] D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," J. Opt. Soc. Am. **56**, 1372-1379 (1966).

[6] A. Labeyrie, "Attainment of diffraction limited resolution in large telescopes by Fourier analyzing speckle patterns in star images," Astron. Astrophys. **6**, 85-87 (1970).

[7] K. T. Knox and B. J. Thompson, "New methods for processing speckle pattern star images," Astrophys. J. Lett. Ed. **182**, L133-L136 (1973).

[8] A. W. Lohmann, G. Weigelt, and B. Wirnitzer, "Speckle masking in astronomy: triple correlation theory and applications," App. Opt. **22**, 4028-4037 (1983).

[9] M. C. Roggemann and B. Welsh, *Imaging Through Turbulence*, (CRC Press, Boca Raton, 1996).

[10] G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," Opt. Lett. **13**, 547-549 (1988).

[11] R. W. Gerchber and W. O. Saxton, "Practical algorithm for the determination of phase from image and diffraction plane pictures," Optik **35**, 237-246 (1972).

[12] B. L. K. Davey, R. G. Lane, and R. H. T. Bates, "Blind deconvolution of noisy complex-valued image", Opt. Commun. **69**, 353-356 (1989).

[13] Y. Yang, N. P. Galatsanos, and H. Stark, "Projection-based blind deconvolution," J. Opt. Soc. Am. A **11**, 2401-2409 (1994).

[14] N. Wang, Y. Chen, Z. Nakao, and S. Tamura, "Parallel-distributed blind deconvolution based on a self-organizing neural network," Appl. Opt. **38**, 4345-4353 (1999).

[15] T. J. Holmes, "Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach," J. Opt. Soc. Am. A **9**, 1052-1061 (1992).

[16] I. Kopriva, D. J. Garrood, and V. Borjanovic, "Single-frame blind image deconvolution by non-negative sparse matrix factorization," Opt. Commun. **266**, 456-464 (2006).

[17] B. C. McCallum, "Blind deconvolution by simulated annealing," Opt. Commun. **75**, 547-549 (1988).

[18] D. G. Sheppard, B. R. Hunt, and M. W. Marcellin, "Iterative multiframe superresolution algorithms for atmospheric-turbulence-degraded imagery," J. Opt. Soc. Am. A **15**, 978-992 (1998).

[19] N. Baba, H. Tomita, and H. Miura, "Iterative reconstruction method in phase-diversity imaging," Appl. Opt. **33**, 4428-4433 (1994).

[20] H. R. Ingleby and D. R. McGaughey, "Parallel multiframe blind deconvolution using wavelength diversity," in *Image Reconstruction from Incomplete Data III*, P. J. Bones, M. A. Fiddy, and R. P. Millane, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **5562**, 58-64 (2004).

[21] T. J. Schulz, "Multiframe blind deconvolution of astronomical images," J. Opt. Soc. Am. A **10**, 1064-1073 (1993).

[22] A. M. Darling, "Blind deconvolution for referenceless speckle imaging," in *Digital Image Synthesis and Inverse Optics*, A. F. Gmitro, P. S. Idell, and I. J. LaHaie, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1351**, 590-599 (1990).

[23] H. Miura and N. Baba, "Extended-object reconstruction with sequential use of the iterative blind deconvolution method," Opt. Commun. **89**, 375-379 (1992).

[24] N. Miura, S. Kuwamura, N. Baba, S. Isobe, and M. Noguchi, "Parallel scheme of the iterative blind deconvolution method for stellar object reconstruction," Appl. Opt. **32**, 6514-6520 (1993).

[25] B. Porat, *Digital Processing of Random Signals -Theory and Methods*, (Prentice-Hall, Englewood Cliffs, 1994), Ch. 3.

[26] G. Harikumar and Y.s Bresler, "Analysis and comparative evaluation of techniques for multichannel blind deconvolution," in *Proceedings of the 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 332-335 (1996).

[27] D. Yellin and B. Friedlander, "Multichannel system identification and deconvolution: performance bounds," IEEE Trans. Sig. Process. 47 1410-1414 (1999).

[28] O. Shavli and E. Weinstein, "Maximum likelihood and lower bounds in system identification with non-Gaussian inputs," IEEE Trans. Inform. Theory 40, 328-339 (1994).

[29] D. J. Lee, M. C. Roggemann, and B. M. Welsh, "Cramér-Rao analysis of phase-diverse wave-front sensing," J. Opt. Soc. Am. A 16, 1005-1015 (1999).

[30] J. J. Dolne and H. B. Schall, "Cramér-Rao bound and phase-diversity blind deconvolution performance versus diversity polynomials," Appl. Opt. 44, 6220-6227 (2005).

[31] C. Y. Chi and C. H. Hsi, "2-D blind deconvolution using Fourier series-based model and higher-order statistics with application to texture synthesis", in *Proceedings of the Ninth IEEE Signal Processing Workshop on Statistical and Array Processing*, pp. 216-219 (1988).

[32] J. L. Jensen, J. A. Jensen, P. F. Stetson, and P. Antonius, "Multi-processor system for real-time deconvolution and flow estimation in medical ultrasound," in *Proc. IEEE Ultrasonics Symposium*, pp. 1197-1200 (1996).

[33] http://www.cray.com/products/xd1/index.html

[34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, (Prentice Hall, Upper Saddle River, 1993).

[35] J. W. Goodman, *Introduction To Fourier Optics, 3rd ed.*, (Roberts and Co., Greenwood Village, Colorado, 2005).

[36] C. L. Matson and A. Haji, "Biased Cramér-Rao lower bound calculations for inequality-constrained estimators," J. Opt. Soc. Am. A 23, 2702-2713 (2006).

[37] C.L. Matson, "Resolution, linear filtering, and positivity," J. Opt. Soc. Am. A 15, 33-41 (2006).

[38] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, (Institute of Physics Publishing, Bristol, 1998).

[39] H. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," IEEE Trans. Image Process. 1, 322-336 (1992).

[40] J. C. Dainty, "Stellar speckle interferometry," in *Laser Speckle and Related Phenomena*, 2nd ed. (Springer-Verlag, Berlin, 1984), Chap. 7.

[41] O. von der Luehe, "Estimating Fried parameter from a time-series of an arbitrary resolved object imaged through atmospheric turbulence," J. Opt. Soc. Am. A 1, 510-519 (1984)

[42] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes*, Cambridge Press, Cambridge (1992).

[43] http://www.inference.phy.cam.ac.uk/mackay/c/macopt.html

[44] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, R. Y. Tsai and T. S. Huang, eds, (JAI Press, 1984), vol. 1, pp. 317-339.

[45] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," IEEE Trans. Image Process. 15, 1413-1428 (2006).

[46] http://www-unix.mcs.anl.gov/dbpp/text/book.html, Section 3.4.

[47] http://www-unix.mcs.anl.gov/mpi/mpich1/

[48] http://en.wikipedia.org/wiki/InfiniBand

[49] http://en.wikipedia.org/wiki/Myrinet

[50] http://www.mhpcc.hpc.mil/

[51] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Ch. 5, Addison-Wesley, Reading (1992).

[52] C. L. Matson and D. W. Tyler, "Primary and secondary superresolution by data inversion," Opt. Express 14, 456-473 (2006). http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-2-456

[53] D. R. Gerwe and P. S. Idell, "Cramér-Rao analysis of orientation estimation: viewing geometry influences on the information conveyed by target features," J. Opt. Soc. Am. A 20, 797-816 (2003).

[54] http://www.maui.afmc.af.mil/

[55] L. C. Roberts, Jr. and C. R. Neyman, "Characterization of the AEOS adaptive optics system," PASP 114, 1260-1266 (2002).